

## Appendix Exhibit C

**The Manville Personal Injury Settlement Trust X-Ray Audit:  
An Assessment of the Identification of the Underlying Disease Process  
Implications for Medical Review by Certified B-readers**

**A. Russell Localio**

**Allen Kunselman**

**Bryan Crissinger**

**Biostatistics Section, Department of Health Evaluation Sciences  
Penn State University College of Medicine  
Hershey, PA 17033**

**Center for Clinical Epidemiology and Biostatistics  
University of Pennsylvania  
Philadelphia, PA 19104**

**February 24, 1998**

**CRMC 0116759**

## Table of Contents

I. Introduction .....	Page 1
II. Background .....	Page 1
A. Background on the Trust .....	Page 2
B. Operation of the Trust and the Categorization of Claims .....	Page 3
C. Background of the Trust's Medical Audit Program .....	Page 3
D. X-Ray Audit Process .....	Page 3
III. Study Objectives .....	Page 4
A. Overview .....	Page 5
B. Primary Objectives .....	Page 5
C. Ancillary Objectives .....	Page 5
IV. Data Collection .....	Page 5
A. Overview .....	Page 6
B. Claims Database .....	Page 6
C. Duration of Exposure File .....	Page 6
D. Rate of Filing of Claims for Mesothelioma File .....	Page 6
E. Trust's Determination of Underlying Disease Process (UDP) .....	Page 6
F. Claimant's B-reader File .....	Page 7
V. General Methodology .....	Page 7
A. Definition of a Dispute over the UDP .....	Page 9
B. Primary Objective Methodology .....	Page 9
1. Initial Analysis .....	Page 9
2. Statistical Modeling of the Predictors of a Dispute .....	Page 10
a. Special Treatment of State and Industry .....	Page 11
b. Identity of Key B-reader .....	Page 11
C. Ancillary Objective Methodology .....	Page 12
1. Modeling the Adjusted Differences Among Law Firms and Claimants' Physicians .....	Page 12
2. Comparing the Medical Audit B-readers .....	Page 13
D. Power of the study .....	Page 13
VI. Findings .....	Page 13
A. Primary Findings .....	Page 14
1. Frequency of Disputes .....	Page 14
2. Predictors of "Disputes" .....	Page 15
a. Model A: Excludes Law Firm and Claimant's Physician .....	Page 15
i. Marginal Dispute Rates for Pre-Audit Claim Category .....	Page 16

b. Models B, C, & D: Assessing Importance of Law Firm and Claimant's Physician .....	Page 17
i. Model B: Assessment of Law Firm .....	Page 18
ii. Model C: Assessment of Claimant's Physician .....	Page 18
iii. Model D: Assessment of Combination of Firm and Physician .....	Page 19
c. Summary .....	Page 20
B. Ancillary Findings .....	
1. Adjusting for Key B-reader to Assess Variation Among Law Firms, Claimants' Physicians, or a Combination of the Two .....	Page 20
a. Assessment of the Variation Among Law Firms Adjusted for Key B-reader .....	Page 21
b. Assessment of the Variation Among Claimants' Physicians Adjusted for Key B-reader .....	Page 22
c. Assessment of the Variation Among Combinations of Law Firms and Claimants' Physicians Adjusted for Key B-reader .....	Page 22
d. Summary .....	Page 23
2. Comparisons Among MA B-readers .....	Page 23
a. Marginal Rates of Finding Disease Process .....	Page 24
b. Head-to-Head Comparisons of the MA B-readers .....	Page 25
c. Summary .....	Page 26
VII. Comments .....	Page 27
Tables 11a - 20c .....	Pages 29 - 48
Appendix A .....	
Table 1 — Categorization Criteria .....	Page 49
Table 2 — Analysis Information .....	Page 53
Table 3 — Ranking of Law Firms By Rate of No Dispute .....	Page 56
Figure 1 — Graphical Ranking of Law Firms By Rate of No Dispute .....	After p.57
Table 4 — Ranking of Physicians By Rate of No Dispute .....	Page 58
Figure 2 — Graphical Ranking of Physicians By Rate of No Dispute .....	After p.59
Table 5 — Ranking of Law Firm-Physician Combinations .....	Page 60
Figure 3 — Graphical Ranking of Law Firm-Physician Combinations .....	After p.62
Table 6 — Adjustment for Maldistribution via Indirect Standardization .....	Page 63
References .....	Page 67

## I Introduction

This report responds to a request by the Manville Personal Injury Settlement Trust (the Trust) to measure the extent of dispute regarding x-ray interpretation of the Underlying Disease Process (the UDP) of the claimants during the Trust's medical (x-ray) audit process. Specifically, we were asked to assess the statistical significance of such disputes and to identify and then measure the relative importance of the various factors associated with such disputes. We were also asked, in the event inter-reader variation was determined to be a statistically significant factor, to determine the relative importance of such inter-reader variation in estimating the rate of dispute between the Trust and any given law firm.

Although such consulting arrangements are not a major part of the work and research activities of the Biostatistics Section of the Department of Health Evaluation Sciences at Penn State University or of the Center for Clinical Epidemiology and Biostatistics at the University of Pennsylvania, the faculty and staff regularly engage in comparable biomedical research projects with government and industry. Our methods represent standard statistical practices and rely on well-accepted statistical tools. The consulting arrangement began in July 1996 and continues.

None of the persons who have worked on this project has any prior relationship to the Trust or to any of the individuals or firms whose names have surfaced during the course of this project. Nor is anyone a claimant or a relative of a claimant before the Trust or any other organization involved in the compensation for exposure to asbestos. We have no financial interest (direct or indirect) that might affect the conduct or reporting of this work. The Trust has offered us continuing support in terms of the provision of data necessary to answer the questions posed; in addition as we have identified issues for further investigation, the Trust has offered continuing advice on the limitations of raw data and on the interpretation of results.

## II. Background

### A. Background on the Trust

The Manville Personal Injury Settlement Trust (the Trust), was established in 1988 to resolve claims against Johns-Manville Corporation arising from its production and sale of asbestos and asbestos products. The purpose of the Trust is to assume and resolve the corporation's obligations for personal injury or death allegedly stemming from exposure to asbestos mined by the corporation or its predecessors, or from exposure to products containing asbestos manufactured or sold by the corporation or its predecessors. The Trust is to use its assets to deliver fair, adequate and equitable compensation to bona fide asbestos health claimants and beneficiaries, whether presently known or unknown, without overpaying or underpaying any claims and with settlement to be preferred over arbitration, arbitration to

be preferred over resort to the tort system, and the fair and efficient resolution of claims to be preferred over all else:

#### B. Operation of the Trust and the Categorization of Claims

Law firms representing claimants submit the evidence to support their claims on the Trust's individual Proof of Claim (POC) forms. These forms are reviewed and the claims are categorized based on the criteria set forth in the Trust Distribution Process (TDP), one of the principal governing documents of the Trust. The TDP and the claim valuation criteria it contains were negotiated beginning in 1992 and agreed to by the Trust and representatives of its beneficiaries in 1994. As set forth in the TDP, the goal of the Trust is to treat all claimants equitably. In furtherance of this goal, the TDP established a Schedule of Asbestos-Related Disease Categories and Values (the Schedule) to enable many claims to be resolved more quickly. Each claimant retains the right to elect individual claim evaluation.

The process for determining the liquidated value of any claim to be paid from the assets of the Trust includes an initial determination of whether the claim meets the Categorization Criteria for one of seven Scheduled Diseases that are listed on the Schedule of Asbestos-Related Disease Categories and Values described in Appendix A, Table 1. The Scheduled Diseases are Bilateral Pleural Disease (Category 1), Nondisabling Bilateral Interstitial Lung Disease (Category 2), Disabling Bilateral Interstitial Lung Disease (Category 3), Other Cancers (Category 4), Lung Cancers One (Category 5), Lung Cancers Two (Category 6), and Malignant Mesothelioma (Category 7). In general, if the claim qualifies for categorization, the claimant will be offered the Scheduled Value for the Scheduled Disease.

If a claimant does not meet the Categorization Criteria for a Scheduled Disease, or the claimant decides to reject the Trust's Scheduled Value offer, and in certain other circumstances, the claimant may elect to have the claim individually evaluated by the Trust. Claimants can have their unresolved disputes over the categorization and valuation of their claims submitted to arbitration. Claimants who lose nonbinding arbitration over the value of their claims may file a suit in the traditional tort system.

When a roof of claim is reached in the FIFO queue, the Trust evaluates the claim to determine whether the claim meets the Categorization Criteria for a Scheduled Disease and advises the claimant of its determination. If a Scheduled Disease is determined, the Trust generally tender to the claimant an offer of payment of the Scheduled Value for the Scheduled Disease. If the Trust determines that a claim does not meet the Categorization Criteria for a Scheduled Disease or if a claimant disagrees with the Scheduled Disease determination made by the Trust, the claimant may dispute the determination. Upon receipt of written advice from the claimant of such a dispute, coupled with the claimant's written statement of the basis for the dispute and any supporting documentation, the Trust shall reevaluate the claim in light of all then-available documentation and advise the claimant of its

determination. If on reevaluation the Trust determines that the claim qualifies for placement in a Scheduled Disease Category or in a different Scheduled Disease Category than the Trust originally determined, the Trust shall tender an offer in the amount of the Scheduled Value for the Scheduled Disease.

If the claimant still disputes the Trust's categorization of the claim or denial of categorization, the claimant may elect arbitration of the categorization or individual evaluation.

#### C. Background of the Trust's Medical Audit Program

The Trust's Medical Audit Program assists in the assessment of the reliability of medical evidence provided by the claimants to substantiate their asbestos-related personal injury claims, including the interpretation of x-rays. Claims are selected at random from the entire claim population eligible for payment, including those claims originally categorized as well as those recategorized based on the submission of new medical information. In general, the Trust only selects for medical audit those claims categorized as Category 1, 2, 3, 4, and 6. However, for special projects, the Trust may select claims that do not meet the categorization criteria for inclusion in Categories 1-7, hereafter referred to as Category 0 claims. In addition, based on its medical audit pass rate, a law firm's entire population of claims eligible for payment may be medically audited. Claims associated with doctors or medical facilities whose reliability is unknown or subject to question may also be identified and subjected to audit.

When a claim is selected for medical audit the law firm receives notice of medical audit, and a request for specific information on the claim, which might include: the original x-ray, pathology tissue samples, a request for the claimant's full medical history, or a requirement for an Independent Medical Examination. A claim can also pass medical audit without x-ray review if the claimant submits an affidavit that he has settled his claim with one or more asbestos co-defendants in an amount equal to twice the Scheduled Value of the acknowledged Scheduled Disease Category or he provides acceptable corroborating medical reports.

#### D. X-Ray Audit Process

In most cases, however, the audit consists of a review of the claimant's x-rays by independent B-readers retained by the Trust (Medical Audit B-reader or MA B-reader). The designation of B-reader denotes the highest certification available for physicians trained in the use of the International Labour Organization (ILO) System for classification of x-rays for the presence of dust-related lung condition known as pneumoconiosis. It requires an applicant to pass the National Institute for Occupational Safety and Health (NIOSH) proficiency examination for classifying chest x-rays for pneumoconiosis (including asbestos-related

conditions) (Wagner 1992). The MA B-readers are requested to complete the standard ILO Sheet for roentgenographic interpretation of the x-rays, which the Trust then records in its claims processing system (CPS). The MA B-reader selects the x-ray film he/she feels is most appropriate from those supplied by the claimant. The results of each reading are compared against the Trust's pre-audit classification of the alleged injury based on the Categorization Criteria described in section II.B and the pre-audit disease process/Category is either confirmed or modified. In general, the Trust has assigned x-ray films in batches of about 30 to the next available MA B-reader in the order in which films were received from the law firms.

If the findings of the MA B-reader to whom the claimant's x-rays are first submitted support at least the level of disease determined by the Trust to be demonstrated by the evidence submitted by the claimant, no further x-ray audit action is taken and the claim is returned to the queue for further processing. If the first MA B-reader's findings are less severe than the pre-audit categorization by the Trust, a second randomly selected MA B-reader (not the initial reviewer) reviews the x-rays. The second MA B-reader is unaware of the results of the first review and that he is the second MA B-reader to review the film. If the second MA B-reader's findings are also less severe, the Trust recategorizes the claim based on the most serious compensable disease supported by the findings of the two MA B-readers. If the second MA B-reader's findings are at least as severe as the pre-audit categorization, the claim is returned to the queue for further processing. If a higher level of disease is found by any MA B-reader than that categorized by the Trust, the Trust recategorizes the claim at the highest value.

Therefore, each claimant has two opportunities to have their claim demonstrate a level of disease at least as severe as the Trust's pre-audit disease categorization. We refer to the first review of a claimant's x-rays by a MA B-reader as the "first stage" review and the second review as the "second stage" review. We also refer to the MA B-reader who arrived at the most severe (highly compensated) diagnosis based on the Schedule and whose finding was the basis for the "final" disease process as the "Key B-reader." We use the term "Key B-reader" in this special context throughout the remainder of this report.

### III. Study Objectives

#### A. Overview

Our analysis had a primary objective and several ancillary objectives. The primary objective was to describe the extent of disputes over the UDP, and attempt to identify the principal predictors of such disputes. If inter-reader variation was determined to be a statistically significant factor, an ancillary objective was to assess the relative importance of such inter-reader variation in estimating the rate of dispute between the Trust and any given law firm or the claimants' physicians. Other ancillary objectives included assessing the

implications of inter-reader variation on the current method of auditing all claims filed from certain law firms (and/or diagnosed by certain physicians) with a high rate of dispute.

#### B. Primary Objectives

Our primary task was to examine the disparities, if any, between the claimant's alleged UDP (as assigned by the Trust before the x-ray audit process) and the UDP that resulted from the reading of the claimant's x-rays by the MA B-readers during x-ray audit. This task included an assessment of the factors that were associated with findings of "dispute" between the alleged UDP and the audit UDP. The task included the use of multivariable statistical models that predict, at least in part, the disputes between claimant and MA B-reader as to the UDP. It also attempted to assess the absolute and relative statistical significance of such predictors.

#### C. Ancillary Objectives

As an outgrowth of having achieved our primary objective, we determined that inter-reader variation was substantial. In addition, we found that different MA B-readers manifested markedly different rates at which they found bilateral interstitial lung disease, bilateral pleural disease, or none of these disease processes. We also determined that claims, although randomly allocated to the Medical Audit B-readers, did not result in the same frequency of assignment of MA B-readers by law firm. For that reason, the identity of the MA Key B-reader became a potential confounder in any analysis of the differences in rates of "dispute" between the Trust and a law firm. Therefore, we were obligated to assess further the extent of this confounding and then to estimate the implications of this variation in the reading of the x-rays by the MA B-readers on the probability of having a dispute and the overall characterization of the quality of a law firm's claims.

### IV. Data Collection

#### A. Overview

We received several computerized databases from the Trust on which to base an analysis. We had no access to paper records. These included data on 6,524 claims that had been released from Medical Audit having undergone the x-ray review by either one or two MA B-readers. These 6,524 claims had been classified as pre-audit categories 0, 1, 2, 3, 4, and 6. Some of the data were copied from the information supplied by the claimants on the Trust's Proof of Claim (POC) Form and provided a general profile of the claimant. Other data elements were based on the Trust's evaluation of the claim and derived from information in the POC Form, updates to the POC Form, or the findings of the x-ray audit. All claims were identified by number to protect the confidentiality of the claimants. Table 2 of Appendix A lists each variable name and definition. In this report we shall always try to

define variables.

#### B. Claims Database

We received several databases from the Trust. The need for additional data became evident following our preliminary assessment of the initial database. This first dataset included computerized data on 6,524 claims that had been released from Medical Audit as of approximately November 30, 1996, having undergone the x-ray review by either one or two MA B-readers. Data were taken largely from the information requested as part of the Trust's POC Form and evaluative information as determined by the Trust. Variables included claim category, the claimant's date of birth, year of exposure, year of filing the claim, the industry, the state, the law firm representing the claimant, and the claimant's physician. In addition, the file reported the results of the x-ray audit: the identity of the MA B-reader and the results of the review, including the radiograph quality, the disease classification, and the "ILO" scores.

#### C. Duration of Exposure File

To these data were added at the claimant level additional information on the duration of the claimant's exposure to asbestos, the first year of the exposure, and the number of jobs held. This file was linked to the original dataset by means of a unique claim number to achieve a better understanding of the possible influence of exposure on the probability of a dispute regarding the UDP.

#### D. Rate of Filing of Claims for Mesothelioma File

This file was created at the Trust to generate a rough proxy for the severity of cases filed across industries and states. Owing to the large number of industries and states as well as the need for a method of adjusting for many factors (see section V.B.2), possibly including these, in several of our analyses, we required from the Trust a simple indicator that might explain differences in typical severity mix of cases by state and industry. The Trust provided data by industry and state on the number of claims filed and the subset of those claims that were filed for mesothelioma. Thus, this file indexed the overall severity of the claims filed by industry and state. Each of the files were linked to the main claims database by industry or state as appropriate. Details on when we used this file appear in the General Methodology section V.B.2.a.

#### E. Trust's Determination of Underlying Disease Process (UDP)

The Trust's staff assigns a category to a claim by responding to a set of questions (core questions) that mirror the criteria set forth in the TDP (Appendix A, Table 1). The Trust provided for each claim a file that listed answers to selected core questions (that

establish disease process).

#### **F. Claimant's B-reader File**

This file consisted of the names and the B-reader status of 134 physicians who have provided diagnoses in support of claimants for whom the medical audit records were available. Of these, 44 were identified as B-readers. Linking this file with the other databases supported our analyses of the impact of B-reader certification on the pass rates of claims. Our initial analyses on differences among the MA B-readers prompted an investigation of any possible association between the B-reader status of the claimants' physicians and the probability of pass by the MA B-readers. We used this file to provide additional guidance on the performance of MA B-readers relative to the opinions of any B-readers used by the claimants. Our conclusions from this analysis appear in the Comments section (VII).

#### **V. General Methodology**

##### **A. Definition of a Dispute over the UDP**

Disputes over the appropriate Category assignment can occur when the Category assigned by the Trust after its review of the evidence supplied by the claimant is more severe (based on the Schedule) than the Category assigned to the claim by the Trust following medical (x-ray) audit. Because the assignment of a Category is a function of not only the strict medical evidence, but also other criteria such as exposure, product identification, causation, and latency, this study focused on disputes over the UDP. While the Category is based on the satisfaction of all the minimum criteria set forth in TDP, the UDP is based solely on the diagnosis of certain relevant disease processes, namely bilateral pleural disease (BPD) or bilateral interstitial lung disease (BILD). The presence of these disease processes is a minimum requirement for a claim to be categorized as displayed in Table 1. For the purposes of this report, Category 0 claims were not considered in the definition of a dispute over the UDP.

Table 1

Correspondence Between Pre-Audit "Category" and the  
Minimum Underlying Disease Process (UDP)

Pre-Audited Category	Minimum Underlying Disease Process
Category 1	BPD
Category 2	BILD*
Category 3	BILD
Category 4	BPD or BILD**
Category 6	BPD or BILD**

\* also requires causation statement or finding of pleural disease

\*\* or pathological evidence

As applied by the Trust, the minimum requirement to establish bilateral pleural disease (BPD) is the presence of bilateral pleural abnormalities consistent with asbestos-related pneumoconiosis, and the corresponding minimum requirement for bilateral interstitial lung disease (BILD) is bilateral parenchymal abnormalities consistent with asbestos-related pneumoconiosis and profusion (ILO score) of 1/0 or greater. Because the Schedule assigns a greater value to non-malignancy claims with bilateral interstitial lung disease than bilateral pleural disease, the UDP for each claim was defined as bilateral interstitial rather than the bilateral pleural where both were present.

The UDP is directly determined by the Trust from the medical evidence supplied by the claimant and presented in the claim file (Claimant's UDP) and can also be determined based on the ILO Classification sheet completed by the Key B-reader from the reading of the x-ray (B-reader's UDP). The comparison between the results can be represented as in Table 2:

Table 2

Cross Classification of Disease Process by Claimant  
and Conclusions of Disease Process by the "Key B-reader"

Claimant's UDP	Key B-reader's UDP		
	None	BPD	BILD
None	Agree	Disagree	Disagree
BPD	Dispute	Agree	Disagree
BILD	Dispute	Dispute	Agree

Since the Trust assigns to the claim the most severe disease process that results from the reading of the x-ray, the only events that create a potential "dispute" with the claimant are those in which both MA B-readers have found a less severe disease process than that alleged by the claimant. That is, Table 2 may be decomposed into a binary response: "dispute" and "no dispute," where "no dispute" is defined as agreement or disagreement. In order to analyze this disparity between claimants' medical reports and the ILOs of the MA B-readers, we examined the factors in the claim (or in the x-ray audit process) that seemed to explain or "predict" whether the process ended in a dispute.

## B. Primary Objective Methodology

### 1. Initial Analysis

The purpose of the initial analysis was, first, to get a sense of the numbers and distributions of claims according to all of the factors in the databases, and second, to assess whether the rates of disputed or not disputed claims, on the issue of disease process, between the claimant's allegations and the Medical Audit's B-reader, differed according to these classifications. We looked for simple associations between each factor, one by one, and the probability of dispute. Statistical significance of these associations was tested using chi-square ( $\chi^2$ ) statistics (SAS PROC FREQ, v 6.12, SAS Institute, Cary, NC). Factors that were highly associated with the outcome (dispute/no dispute) were considered as potential predictors for a multivariable assessment of dispute (see section V.B.2).

### 2. Statistical Modeling of the Predictors of a Dispute

A statistical model's principal purpose is to determine what factors explain an outcome. In the present case, the outcome of interest is whether the UDP is disputed during

Report, Manville Personal Injury Settlement Trust x-ray Audit 12/09/97  
Dept of Health Evaluation Sciences, Penn State University College of Medicine

CRMC 0116770

the x-ray review of the Medical Audit. This binary outcome (no dispute/dispute) is appropriately modeled using multivariable, logistic regression (Armstrong 1994). Since the outcome is binary (dispute/no dispute), modeling dispute as the outcome is the equivalent of modeling no dispute as the outcome; the chi-square statistics for the predictors remain unchanged. A regression model tries to examine the influence of each factor, while controlling for the other factors in the model. It asks the question whether factor A demonstrates an association with the outcome when another potential factor (B) is held constant. In this manner, multivariable regression allows one to view several factors simultaneously and identify the ones that in lay terms are the most "important" determinants of outcome. The regression process involves the "adjustment" or "controlling" for multiple factors when focusing on one factor alone.

We obtained from our Initial Analysis (see section V.B.1) the factors (a subset of all available variables) that might be candidate, i.e., possible, predictors of a dispute. Our analysis was limited to available data. A common problem in any analysis is the potential influence of omitted factors. They are omitted either because the data do not exist, or because data cannot, at a reasonable cost, be collected and used. The challenge of omitted factors applies to all statistical analyses.

Initially, we divided factors such as age into groupings that resulted in moderate to large numbers of claimants in each group. We did the same for the factors related to the claimant's length of exposure: duration of exposure and first year of exposure. They were categorized to give sufficient numbers of observations per group. The year the claim was filed was included as a variable to allow us to detect whether claim dispute rates might be changing over time. As with the other factors, this variable was also categorized to allow for nonlinear relationships between the factor and the outcome of dispute/no dispute. (By nonlinear relationships we mean that the dispute rate might rise or fall irregularly from the early to the late filed claims). Model building proceeded with a backward selection process. That is, starting with a model with all possible candidate factors obtained from the Initial Analysis (section V.B.1), we dropped those that demonstrated very little explanatory power.

#### a. Special Treatment of State and Industry

State and industry presented special problems. Owing to the large number of states and industries represented in the datasets relative to the number of claims, we could not determine the importance of industry and state directly, while we also attempted to assess the importance of other multi-level factors, such as law firm and claimant's physician. That impediment led us to request from the Trust some additional information about these potential factors.

The Trust provided a dataset (see section IV.D) on the rate by state and industry of the filings of claims for mesothelioma. These rates represented one measure of the severity level

of claims and were used only for preliminary analyses in which we sought to identify the most important predictors of a dispute. We hypothesized that states and industries with a history of filing greater relative numbers of mesothelioma claims might have a higher probability of having claims not disputed in the x-ray audit. (As the Trust does not audit claims based on mesothelioma, there were no mesothelioma claims in this dataset of 6,524).

The relationship between these state and industry measures of claim severity and the probability of dispute was tested, as with all other factors, via multivariable regression analysis. By using these summary measures, we then were able to include law firm (and claimant's physician) directly in the later analyses, having eliminated some of the less important factors. The enormous computing requirements of our assigned task was the principal reason for use of these mesothelioma rates.

The stability of the claim severity rates using the filings of mesothelioma cases as a proxy for severity depends on the number of claims filed. To arrive at stable estimates of the true rate of filing, therefore, we used empirical Bayes methods (Casella 1985). These methods anticipate regression to the mean. Each rate was shrunk toward the overall mean rate by an amount that depended on how far the individual (state or industry) was from the mean and how many claims have been submitted (from the state or industry). The smaller the number, the greater the degree of shrinkage. Theory and empirical research suggest that the true rate will be more accurately reflected in the empirical Bayes (shrunk) estimates than in the raw rates.

Empirical Bayes methods were implemented by generalized linear mixed models (Wolfinger 1993) in which the outcome is binary (mesothelioma case or not), the intercept (overall mean) was the sole fixed effect, and the state or industry were random effects. We fit two separate models (state and industry) (SAS PROC MIXED, v 6.12, SAS Institute, Cary, NC). The estimates of the random effects then produced empirical Bayes shrunk estimates of the true rates of filing mesothelioma claims by state and industry. These shrunk estimates were then used in subsequent multivariable, logistic regression models when the need arose to control for state or industry, with each of these factors entering the model as a linear term.

## 2. Identity of Key B-reader

Based on our early regression modeling of dispute rates, we concluded that an important predictor was the identity of the Key B-reader. Using a regression method we were able to assess the relative importance of the identity of the Key B-readers after controlling for other characteristics of the claims. We therefore could ask whether the observed differences among the MA B-readers could be explained by differences in the claims they reviewed or whether the differences among MA B-readers persisted in spite of the possibility of differences in their claims.

Report, Manville Personal Injury Settlement Trust x-ray Audit 12/09/97  
Dept of Health Evaluation Sciences, Penn State University College of Medicine  
Center for Clinical Epidemiology and Biostatistics, Univ of Pennsylvania

CRMC 0116772

### C. Ancillary Objective Methodology

#### 1. Modeling the Adjusted Differences Among Law Firms and Claimants' Physicians

After estimating the factors that predicted a dispute, we then estimated whether claims with some law firms or claimants' physicians were more likely to end in dispute. The goal of this second set of statistical models was to assess whether differences in the propensities of the Key B-readers to find disease process (see section V.C.2), could explain the apparent differences among the law firms (and claimants' physicians) in their findings of disease process.

For this task we resorted to a generalized linear mixed model. In these models, the "fixed effect" to be controlled for or adjusted for was the Key B-reader. This adjustment was especially important for two reasons: (1) although claims are allocated to B-reader in a random process, the distribution of the Key B-readers in the cases in this sample were not completely random across firms; (2) the MA B-readers differ, in some cases markedly, in their propensity to find disease. Therefore, the Key B-reader was what is termed a "confounding" factor in any analysis of the differences in rates of dispute by law firm (or physician). Treating law firms (or claimants' physicians) as random effects in the same regression model (with Key B-reader) allowed us to analyze the variation among the law firms (or claimants' physicians), but because of the small number of claims relative to the number of law firms and physicians, we had to employ a different statistical model for this analysis. As with the empirical Bayes models discussed in section V.B.2.a, the use of random effects in generalized linear mixed models compensated for the small numbers of claims at the law firm or claimant's physician level by shrinking the observed rates of dispute toward the overall mean.

The statistical method employed for this analysis attempted to assess from past data what might be the relative performance of the law firms and physicians. Some of the classic work in this field studied batting averages of major league baseball players (Efron 1975). It is well known that early on in the season when the number of at bats is low, some hitters have unusually high or low batting averages. As the number of at bats increase over the season, averages tend to settle down with both the high and low rates tending to move towards the overall mean of all players. It would be unwise to predict the true batting average based on only a few at bats. Likewise, assessing whether certain law firms or physicians have high or low dispute rates necessitates special methods that consider not only the observed dispute rate but also the number of claims and the overall mean dispute rate.

#### 2. Comparing the Medical Audit B-readers

Comparisons were made among the MA B-readers with respect to their classifications

of claims by the underlying disease process. To assess the degree of agreement among MA B-readers we focused on two dimensions of agreement. First, to achieve high levels of agreement, the MA B-readers would need to exhibit similar UDP rates on the x-rays they reviewed. We examined MA B-reader-specific rates of finding disease process on both their first stage and second stage reviews. To obtain more stable agreement rates, MA B-readers that reviewed only a small number of claims were excluded from B-reader-specific analyses. This exclusion criteria reduced the number of B-readers from seven to five. The MA B-readers are randomly denoted as MA B-readers A, B, C, D, and E to preserve confidentiality.

Additionally, we examined all possible pairs of MA B-readers to assess how often they agreed. This analysis cross-classified the responses by MA B-reader pairs to permit a display of the patterns of agreement between the pairs. The sample sizes were dependent on the number of claims that were reviewed twice by each MA B-reader pairing that examined the same set of x-rays. With five MA B-readers there are 10 possible pairings, each of which was analyzed for the type of disease found.

#### D. Power of the study

The ability of an analysis to detect from a sample the true differences in rates in the population is called statistical power. Generally, statistical power increases as the sample size increases. In the present situation, the sample of claims was large in total. It was relatively large for many of the groups or subgroups of claims as well.

The power to detect differences in the dispute rates depends on the sizes of the subgroups for comparison. For example, the five MA B-readers analyzed in this study performed hundreds of B-readings, at either the first or second stages of the x-ray audit. In the case of two groups of 500 claims each, this analysis would have a 90% chance of being able to find a statistically significant difference in the two rates of dispute, if in fact the true dispute rates are as close as 50% and 60% (10 percentage points difference). Thus, for many of the comparisons that we anticipated making, there was more than adequate statistical power to detect differences in the rate of disputes of relatively modest size. For other comparisons, for example a comparison between one law firm and another, the number of claims per law firm was often inadequate to support a statistical analysis.

### VI. Findings

#### A. Primary Findings

As stated in section III.B, the primary objective was to examine the disparities, if any, between the claimant's alleged UDP (as assigned by the Trust prior to x-ray audit) and the UDP assigned by the Key B-reader after review the claimant's x-rays.

Report: Manville Personal Injury Settlement Trust x-ray Audit 12/09/97  
Dept of Health Evaluation Sciences, Penn State University College of Medicine  
Cranston, PA 19022

CRMC 0116774

## 1. Frequency of Disputes

The frequency of "disputes" in the UDP between the claimant and the Key B-readers appears in Table 3.

Table 3

Finding of Underlying Disease Process by Key B-reader

Rates of Dispute and No Dispute (N=6,482 claims)

Key B-readers' Findings: Underlying Disease Process				
		None n (%)	BPD n (%)	BILD n (%)
Underlying Disease Process Assigned by Trust	None	4 (0.1)	60 (1.0)	29 (0.4)
	BPD	336 (5.2)	1378 (21.3)	307 (4.7)
	BILD	867 (13.4)	1481 (22.8)	2020 (31.2)

Note: Cells (and their percentages of the total) representing disputed cases are in bold.

This table includes all claims except for: (1) those in which the claimant's underlying disease process is unknown from the core questions (36 claims) and (2) six cases for which the Trust could not match the audited category to the responses to the core questions, i.e.,  $6524 - 36 - 6 = 6482$ . Core questions refers to a set of questions that the Trust answers with respect to each claim that determines whether, according to the Schedule, that claim meets the minimum criteria for categorization as a Scheduled Disease.

From this table the reader can see that 2,684 (the sum of the frequencies in bold) of the 6,482 claims resulted in disputes. Most of the disputes resulted from a claimant's allegation of bilateral interstitial lung disease and a finding of either bilateral pleural (1,481) or no disease process (867) by the Key B-readers. These two categories of dispute represent 36% of all cases. In sum, fifty-nine percent (59%) of all cases were not disputed.

## 2. Predictors of "Disputes"

To ascertain the predictors of disputes, we examined the rates of dispute by several

Report: Manville Personal Injury Settlement Trust x-ray Audit 12/09/97  
 Prepared by: [redacted] From the University College of Medicine

CRMC 0116775

factors, both factors that describe the claim and claimant and factors that describe the Trust's x-ray audit process. For this analysis the outcome (to be predicted) was the presence of a dispute. Based on the results of the Initial Analysis (section V.B.1), candidate predictors included: (1) the characteristics of the claimants — age, the first year of exposure, duration of exposure, number of jobs held, (2) the timing of the claim — year in which a claim was filed against the Trust (or against another defendant) called the FIFO year, (3) the place of exposure (state or industry), (4) the Trust's audit process — category assigned to the claim, identity of the Key B-reader, and (5) the source and quality of the information in support of the claim — claimant's law firm, physician, and the film quality.

Multivariable, logistic regression (section V.B.2) was used to model "dispute" using these candidate variables. A number of these predictors were so highly correlated that we had to reduce the list to a few predictors from each of the five groups listed previously, e.g., only age and the first year of exposure were used from the characteristics of the claimant information. Of most importance, because of the number of states and industries, we had to reduce the state and industry to single predictor variables (as we described in section V.B.2.a). In one sense, the greater the chi-square statistic of a factor in a regression model, the greater the statistical importance of the factor as a predictor of dispute (the size of the estimate of the predictor is also important), but these estimates are less easily compared than are the chi-square statistics. (Overall importance must also be measured by examining the size of the influence of a factor on the probability or odds of a dispute).

For the multivariable assessment of predictors of dispute, certain claims were not considered in the analyses. First, in addition to the claim exclusions noted for Table 3, these analyses required that we drop all cases in which there was a missing value for one of the other candidate predictors in order for the statistical software to model the data appropriately. Second, for these analyses, we excluded two of the MA B-readers who had completed 5 or fewer reviews each. Hence, five MA B-readers remained for analysis. Finally, we eliminated all claims in Category 0 (because they are not routinely audited, although they were present and audited in our working dataset). The analysis was therefore limited to claims in Categories 1, 2, 3, 4, and 6.

#### a Model A: Excludes Law Firm and Claimant's Physician

The first model analyzed does not include the claimant's law firm or physician as potential predictors, because a combination of the very large number of firms (and physicians) relative to the number of claims generated a statistical problem that vastly surpassed the computing capabilities of even our well-endowed computing facilities. Complex statistical models with factors that have many levels require very large datasets for complete analysis of all potential relationships between factors and outcomes, e.g., dispute/no dispute. In section VI.A.2.b, we attempted to assess these individual factors by greatly simplifying the analysis, that is, by examining law firm and physician separately.

Table 4 presents the statistically significant predictors in terms of their relative order of statistical importance in a multivariable regression model, with the chi-square statistic indicating the degree of statistical importance of the factor.

Table 4

## Key Predictors of Dispute in Disease Process

Factor	Degrees of Freedom (df)	Chi-square Statistic ( $\chi^2$ )
Category	4	856
Identity of Trust's Key B-reader	4	583
FIFO year	3	63
First year of exposure	6	29
Claimant's age at diagnosis	4	20
State of claimant's residence	1	13
Film quality	2	10
Industry of claimant's employment	1	2

In Table 4 and subsequent tables, "df" refers to the degrees of freedom in the factor or predictor, and equals the number of levels of the predictor minus 1. The same chi-square statistic is more significant as df becomes lower. Thus one cannot compare chi-square statistics alone. We could not compare the relative statistical significance of the factors with large chi-square values (but with different degrees of freedom) because their statistical significance was so great that making such distinction exceeded the ability of our computers to compute the necessary comparative measures.

However, Table 4 clearly demonstrates that claim category and the identity of the Key B-reader were overwhelmingly the most important predictors and were extremely significant statistically. However, this analysis does not preclude the possibility that law firm and/or claimant's physician are also important factors in predicting the outcome (dispute on the UDP).

## L Marginal Dispute Rates for Pre-Audit Claim Category

As Table 4 depicts, the pre-audit category classification is a very strong predictor of dispute on the underlying disease process. We examined the marginal dispute rates for Category (Table 5), but not controlling (or adjusting) for other factors as was done in Table 4. Table 5 demonstrates that the rate of dispute differs over the categories ( $\chi^2 = 772$ , df=4).

$p < 0.001$ ). In short, the rate of dispute for Categories 2 and 3 was far higher than for the other three categories.

Table 5

Association Between Claim Category and Rate of Dispute/No Dispute (N=6,365)<sup>a</sup>

		No Dispute n (row %)	Dispute n (row %)	Total
Pre-audit Category of Claim	1	1642 (83)	337 (17)	1979
	2	1096 (47)	1212 (53)	2308
	3	878 (44)	1098 (56)	1976
	4	21 (66)	11 (34)	32
	6	54 (77)	16 (23)	70

<sup>a</sup>Category 0 cases were excluded in addition to exclusion criteria listed for Table 3.

<sup>b</sup>These rates of dispute overstate the actual failure rates for cancer claims since any asbestos-related UDP or pathological evidence of asbestos may satisfy the criteria for categories 4 and 6.

The multivariable analyses, an example of which is Table 4, model the simultaneous influence of several factors, including claim Category, to determine the influence of each factor, while controlling for the other factors in the analysis.

## II. Impact of Key B-reader

The impact of the Key B-reader will be discussed in subsequent sections of this report.

### b. Models B, C, & D: Assessing Importance of Law Firm and Claimant's Physician

To address the problem of not having enough data (and having too many firms and/or physicians) as described in VLA.2.a, we reduced the problem to two simpler models. We chose to model separately the impact of law firm and the impact of the claimant's physician on the dispute/no dispute outcome. We could not model both law firm and physician together (and assess their relative importance). These factors are partially "nested" because physicians

are hired by law firms to review groups of cases. The factors are not completely nested, however, since a single physician may be hired by more than one law firm. Our approach was dictated by the limited number of observations relative to the numbers of law firms and physicians as well as the huge computing requirements of the task. Because they were extremely "important" predictors of dispute, we included Category and Key B-reader in each of the following models. To simplify the exercise, we eliminated other predictors determined to be relatively unimportant when compared to Category and Key B-reader. Through this analysis, we attempted to demonstrate in simple but rough terms the relative "importance" of law firm and physician in the prediction of dispute.

#### i. Model B: Assessment of Law Firm

This statistical model added law firm as another factor, but we were forced to limit the number of law firms to those which had filed at least 20 claims (there were 70 such law firms). Using the same method of analysis as in section VLA.2.a, we found that law firm was a strikingly significant predictor of a dispute (Table 6).

Table 6

Relative Importance of Category, Key B-reader, and Law Firm

Factor	df	$\chi^2$ Statistic
Category	4	623
Identity of MA Key B-reader	4	567
Law Firm	69	383

In this model, Category and Key B-reader remain more statistically significant than law firm (because they have much higher chi-square statistics and much lower degrees of freedom). But law firm remains unmistakably an extremely-important predictor. In this range of values of chi-square statistics, every factor, including law firm, exhibits huge levels of statistical significance.

#### ii. Model C: Assessment of Claimant's Physician

We then repeated the analysis of section VLA.2.b.i for claimant's physician and found similar results (Table 7). Again, we limited the analysis to those physicians with at least 20 claims (there were 50 such physicians).

Table 7

Relative Importance of Category, Key B-reader, and Physician

Factor	df	$\chi^2$ Statistic
Identity of MA Key B-reader	4	508
Category	4	394
Claimant's Physician	49	397

Again, Category and Key B-reader remain most significant. Although not quite so important (owing to the large number of physicians and thus a large degrees of freedom), the identity of the claimant's physician remains statistically significant and highly so. This analysis did not consider, however, that physicians might exhibit different patterns when working as an expert for one firm than for another. We had to consider, therefore, the combination of law firm and physician as a potential factor to explain outcome.

### iii. Model D: Assessment of Combination of Firm and Physician

To examine the effect of a combination of law firm and physician, we again had to limit the dataset, this time to those cases in which each combination of law firm and physician had at least 10 claims. (Had we the luxury of more data, we would have pegged this lower bound at 20). Thus, controlling (adjusting) for the category of the claim and the identity of the Key B-reader, we assessed whether there remained any residual explanatory power by the physician when the physician was working for a particular law firm. With the restriction on the minimum frequency of claims, required in order to allow the computer and software to generate an estimate, our analysis was limited to 120 combinations of firm and physician. We found that the combination of law firm and physician is a very significant predictor of dispute, after controlling for the claim category and the Key B-reader (Table 8). But with so few data relative to the possible combinations of law firms and physicians, the task of determining which law firms or physicians, singly or collectively, exhibited the strongest association with dispute required a different mode of analysis. The relationships are complex, owing to the fact that law firm and physician are not independent factors; there appears to be in some cases an interaction between the law firm and the physician. These interactions cannot be further clarified without vastly larger numbers of claims for analysis.

Table 8

Relative Importance of Key B-reader, Category, and Law Firm-Physician Combination

Factor	df	$\chi^2$ Statistic
Identity of Key B-reader	4	464
Category	4	334
Claimant's Law Firm-Physician	119	519

## c. Summary

In summary four factors appear to predict the chance of a dispute as to the UDP far better than any other factor or factors, individually or combined. These factors are claim Category, the identity of the Key B-reader, the claimant's law firm, and the claimant's physician (as well as the claimant's physician when working within one or more law firms). But because of (1) the large number of law firms and physicians, (2) the relatively small number of observations (relative to the number of factors and levels of those factors), and (3) the evidence of substantial possibilities of interactions between law firm and physician, we could not model all of the statistically significant factors simultaneously. For that reason, we cannot make definitive, consistent statements about whether any particular factor is always the single most important factor in explaining whether a claim results in a dispute. We can state, however, that Category, Key B-reader, law firm, and physician (or some combination of law firm and physician) are all individually and collectively strongly associated with dispute in all of our analyses. The strength of the association is manifested both by the extremely high levels of statistical significance of these factors, as well as by the size of the effect of these factors on the probability of a dispute. We call these two dimensions of importance "statistical" and "clinical" significance. Both are present in these factors to a very high degree.

## B. Ancillary Findings

## 1. Adjusting for Key B-reader to Assess Variation Among Law Firms, Claimants' Physicians, or a Combination of the Two

The analysis in Section VLA did not seek to determine which individual law firms or claimants' physicians (or combinations thereof) might be more or less likely to dispute the findings of the Key B-reader. Therefore, the next set of analyses sought to examine further the source of variation, resulting in such large chi-square statistics in Tables 6, 7, and 8, caused by law firm (and physician or the combination of the two). By this analysis we

sought to determine whether and to what extent some law firms and physicians (or combinations of the two) have significantly higher or lower rates of dispute over disease process.

Since Key B-reader was a significant predictor of dispute and the assignment of claims to MA B-readers was not completely random, we could conclude that the Key B-reader was a confounder in any analysis of the differences in the rates of dispute among law firms (or law firm and physicians). To control for this confounder (Key B-reader), we used a generalized linear mixed model (see section V.C.1) using the Key B-reader as the fixed effect and law firm, physician, or law firm-physician combination as the random effect. This model allowed us to arrive at a set of rankings of law firm, physician, and law firm-physician combinations which were adjusted for this highly significant factor (Key B-reader) that is entirely outside the control of the law firm and the claimant's physician. By this method we computed adjusted rankings of law firms by the rate of dispute, compared these adjusted ranks with the raw ranks, and thus determined the influence of the assignment of MA B-readers on any rankings. This analysis allowed us to determine whether any individual law firm might have been audited (or be audited) unfairly because of the large differences among the MA B-readers in their judgments on the presence or absence of an underlying disease process.

We must emphasize, however, that this method of analysis has limited power to identify law firms or physicians (or combinations thereof) with very low numbers of claims as being different from average in terms of their rates of dispute on disease process. Thus, the analyses were limited to law firms, physicians, or a combination of the two which submitted at least 20 claims for the law firms and the physicians, and at least 10 claims for the combination. This number is still below what one would want for adequate statistical power (40 to 50 would be a more appropriate absolute minimum, and 80 to 100 would be preferable).

#### a. Assessment of the Variation Among Law Firms Adjusted for Key B-reader

As set forth in Table 6, we concluded that the variation among the law firms was highly predictive of a dispute. By adjusting for the Key B-reader, we could examine the influence of this factor on the relative rankings of the law firms (in their rates of dispute), and demonstrate the number of law firms that are statistically different from average. The results of this analysis are depicted in Table 3 and Figure 1 of Appendix A. Although we report the raw rates of no dispute for the 70 law firms having 20 or more claims filed, we do not report an "adjusted rate" because of the difficulty of interpreting adjusted rates using our methods of mixed effects statistical models. We do report the adjusted rank, which is a more robust measure of the relative positions of law firms (after adjusting for Key B-reader). We adjusted for Key B-reader in this analysis to be able to estimate what might be the rankings if all MA B-readers were equally calibrated in their judgments about x-ray films.

If the sample size is sufficient, a simple and reasonably accurate method of arriving at an adjusted rate for a single law firm appears as Appendix A, Table 6 (and accompanying text). The method applies only with one of two factors, each with few levels. This situation might apply to MA B-reader, for example, where there were in effect only five levels. The adjustment (for the Key B-reader), called indirect standardization, allows one to determine how well an adjusted rate might correct a raw rate of dispute or no dispute. This method is not appropriate, however, in the comparison of law firms to each other. It also does not replace the multivariable statistical model when the number of factors (and their levels) is not small. Firms with relatively large numbers of claims fared about the same whether one looks at raw or adjusted rates of dispute. Therefore, any accidental departure from complete randomness in the assignment of cases to the MA B-readers did not appear to affect substantially the relative dispute rates of firms.

**b. Assessment of the Variation Among Claimants' Physicians Adjusted for Key B-reader**

From Table 7, we concluded that the variation among the claimants' physicians was also highly associated with a dispute (or no dispute) on disease process. Adjusting for the Key B-reader, we examined the influence of this factor on the relative rankings of the 50 physicians (in their dispute rates) having 20 or more claims submitted. Physicians with relatively large numbers of claims fared about the same whether one looks at raw or adjusted rates of dispute (Table 4 and Figure 2 of Appendix A). Therefore, any accidental departure from complete randomness in the assignment of cases to the MA B-readers did not appear to have affected substantially the relative dispute rates of physicians in this sample of claims.

**c. Assessment of the Variation Among Combinations of Law Firms and Claimants' Physicians Adjusted for Key B-reader**

We also concluded that a combination of law firm and physician was also highly associated with a dispute on disease process (see Table 8). Adjusting for the Key B-reader, we examined the influence of this factor on the relative rankings of the 120 law firm-physician combinations (in their dispute rates) having 10 or more claims submitted.

This analysis revealed 25 law firm-physician combinations in which the rates of dispute were significantly higher than average, after adjusting for the differences in the Key B-readers. The findings appear in Table 5 and Figure 3 of Appendix A. The table provides raw rates, raw ranks, and adjusted ranks as well as the number of cases per firm-physician combination. Adjusted rates are not reported owing to the difficulty of interpreting an adjusted rate produced by our statistical model as previously discussed. For most law firm-physician combinations, the rankings in terms of the order of rates of dispute did not change substantially after adjusting for Key B-reader. For a few law firm-physician combinations, however, the rankings did change to some degree. For example, the

combination of firm #108 and physician #471 (Table 5 and Figure 3, Appendix A) moved from a raw rank of 68 to an adjusted rank of 38. This change, although one of 30 rankings, still did not move the combination into or out of a range that is statistically different from average.

Another interesting finding was that the adjusted ranking of physician depended in part upon the law firm for whom the physician was working. In some cases the physician working for one law firm had a high rate of dispute, while when working for another law firm, the same physician had a much more average rate of dispute (see Appendix A, Table 5).

#### d. Summary

In summary, those law firms, physicians, and law firm-physician combinations with relatively large numbers of claims fared about the same whether one looks at raw or adjusted rates of dispute. Therefore, any accidental departure from complete randomness in the assignment of cases to the MA B-readers did not appear to have affected substantially the relative dispute rates of law firms, physicians, and law firm-physician combinations in this set of claims.

Also, these findings suggest that law firm and physician (and the physician within the law firm) provided a substantial source of the reason why certain disputes on disease process arose. What limited our ability to assess which law firms (or law firm-physician combinations) resulted in the highest (or lowest) dispute rates was the relatively small number of observations per law firm, physician, and law firm-physician combination. The tools of statistics, we have found, are not up to the task of this problem. Ideally, if we had at least 10 disputes (for a law firm-physician combination with a low rate of disputes) and 10 agreements (for a law firm-physician combination with a high rate of disputes), we could use statistical tools with greater success. Statistical methods then might be able to identify with greater power the true outliers, either in terms of high or low rates of dispute over disease process.

## 2. Comparisons Among MA B-readers

In this analysis we included all available data on five MA B-readers. These data included all available categories of claims, 0 through 7. It also included the six questionable cases in which the MA B-reader's disease process ascertainment was inconsistent. As before, however, we excluded the two MA B-readers with very small numbers of reviews (a total of under 10). Thus, we retained five MA B-readers, denoted A through E to preserve confidentiality. Table 9 reports the frequency of reviews by the MA B-readers.

Table 9

## Frequency of x-ray Reviews by MA B-reader

B-reader <sup>a</sup>	Number of First Reviews	Number of Second Reviews	Total
A	876	476	1352
B	1643	1167	2810
C	357	296	653
D	2112	819	2931
E	1530	973	2503

<sup>a</sup>Note: Two MA B-readers with very low numbers of reviews, not used in our analyses, are omitted

#### a. Marginal Rates of Finding Disease Process

For B-readers to agree on the UDP of a given set of patients, each B-reader must be similarly "calibrated." By calibration we mean that a physician must identify (in a given set of patients) underlying disease processes at the same rate. Thus, for example, if the true rate of some underlying disease is 50%, and the B-readers looking at the same group of persons find disease rates of 20% to 80%, then it will be impossible for these B-readers ever to agree on the rate of some underlying disease process. This is a well-known axiom in the field of physician decision-making on patient diagnosis (Feinstein 1985, Elmore et al 1992).

In the present study we examined these marginal rates of finding disease process (Table 10) and found differences that were both clinically and statistically significant ( $\chi^2 = 1084.7$ ,  $df=4$ ,  $p < 0.001$ ). The rates of finding no underlying disease range from 26% to 79% among these five MA B-readers.

Table 10.

Marginal rates of underlying disease process (UDP) as defined by B-readers regardless of whether a first or second stage evaluation.

MA B-reader	MA B-reader Defined UDP			Total n
	None n (row %)	BPD n (row %)	BILD n (row %)	
A	719 (53.2)	490 (36.2)	143 (10.6)	1352
B	751 (26.7)	815 (29.0)	1244 (44.3)	2810
C	514 (78.7)	99 (15.2)	40 (6.1)	653
D	781 (26.7)	2022 (69.0)	128 (4.4)	2931
E	1092 (43.6)	595 (23.8)	816 (32.6)	2503

This range is quite large, even in light of the well-known fact that physicians frequently lack good "calibration" in their diagnostic findings (Felson et al 1973, Kopans 1994, Parker et al 1994, Wagner et al 1993), because the MA B-readers in this analysis have reviewed such large numbers of x-ray films from asbestos claimants. In addition, because these rates are based on very large samples of cases, we are confident that the differences in rates do not result from random variation.

Among the five MA B-readers the rates of finding bilateral pleural versus bilateral interstitial disease also varied widely, even in the case of the two MA B-readers who found no disease at the same rate (27%). One of these found bilateral interstitial disease in 44% of the cases, while the other found this disease in only 4%.

#### b. Head-to-Head Comparisons of the MA B-readers

These findings did not consider that the MA B-readers examined different cases. To assess whether the degree of disparities persisted when B-readers examined exactly the same cases, we performed a more refined set of analyses reported in Tables 11 to 20. We limited the analysis only to the claims which had two B-readings, for we wanted to compare B-readers who examined exactly the same x-ray films. For these cases we had to exclude another group of cases in which there was evidence, by reason of x-ray date, that the MA B-readers might have read films of the same patient but taken on different dates. For this analysis, we wanted to ensure that the pairs of MA B-readers had read the same films. These tables are then based on 3,238 claims.

Report: Marvile Personal Injury Settlement Trust x-ray Audit 12/09/97  
Dept of Health Evaluation Sciences, Penn State University College of Medicine

CRMC 0116786